The charm and beauty of probing CP violation with the CMS detector

Alberto Bragagnolo (University & INFN PD), on behalf of the CMS Collaboration

CERN LPCC EP-LHC Seminar Series
16 April 2024

Introduction

Why we need CP violation?

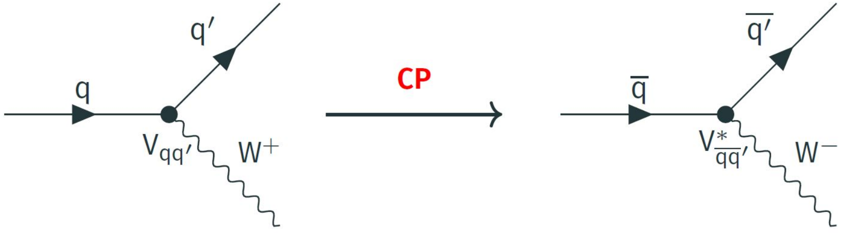
- **Baryon asymmetry remains one of the great mysteries of modern physics**
- Half a century ago, Andrei **Sakharov** proposed three necessary **conditions** for a baryon-generating process:
  1. Baryon number violation
  2. C and **CP violation (CPV)**
  3. Non thermal equilibrium
- In the Standard Model (SM) CP is **conserved** **by the strong and electromagnetic interactions**, but it is **violated** **by the weak force**
  - CPV was first observed in 1964 by Fitch and Cronin using neutral kaons
    - P violation was proposed by Lee and Yang (1956) and experimentally observed by Wu (1957)
- CPV is **allowed** in the SM, but the amount is **insufficient** to account for the observed baryon asymmetry of the universe
  - Sources of CPV beyond the SM have to exist
  - CPV observables are often precisely predicted and very sensitive to new physics

CP violation in the SM

- In the SM quark transitions are possible through flavor-changing weak interactions
- **Information about the strength of the transition is contained in the Cabibbo-Kobayashi-Maskawa (CKM) matrix** 
  - Parameters: 3 angles + 1 complex phase
- **The single complex phase allows for CP violation**
- In the SM, the CKM matrix is **unitary**
  - Unitary conditions can be represented by “unitary triangles”

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

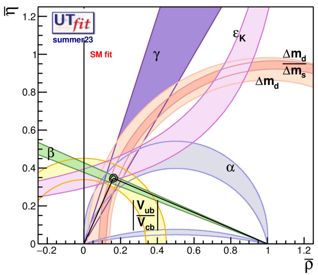
Flavour **CKM** **Mass**



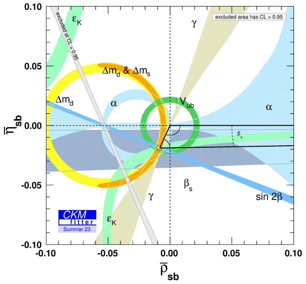
E.g.: u ➔ b amplitude ~ $|V_{ub}|\exp(-i\delta_{CP})$

E.g.: $\bar{u}$ ➔ $\bar{b}$ amplitude ~ $|V_{ub}|\exp(+i\delta_{CP})$

B^0 unitary triangle (UTfit)



B_s unitary triangle (CKMfitter)



Interlude: flavor mixing

- **Neutral K, D, and B mesons are subject to flavor mixing, that is oscillations between their C-conjugate states before decay**
- They propagate as light and heavy mass eigenstates which are described by a superposition of flavor states:

$$|M^0_{L,H}\rangle = p|M^0\rangle \pm q|\overline{M}^0\rangle \qquad (|q|^2+|p|^2 = 1)$$

- The system is characterized by the parameters

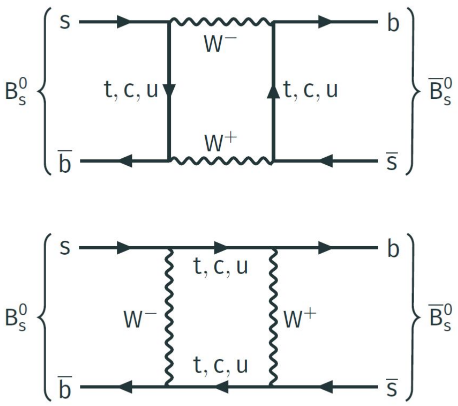
$$m \equiv \frac{m_H + m_L}{2} \qquad \Gamma \equiv \frac{\Gamma_H + \Gamma_L}{2}$$

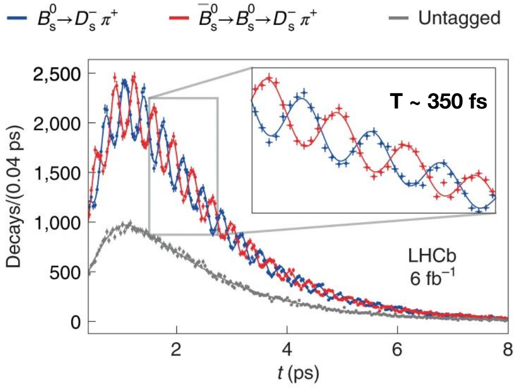
$$\Delta m \equiv m_H - m_L \qquad \Delta\Gamma \equiv \Gamma_L - \Gamma_H$$

Definition may vary

  - The flavor eigenstates oscillate with period T = 2π/Δm
- **CPV in mixing**, i.e. $Pr(M^0 \to \overline{M}^0) \neq Pr(\overline{M}^0 \to M^0)$, **implies |q/p| ≠ 1**

Example: B_s meson mixing





Ref: LHCb Collab. Nat.Phys.18(2022)1-5

Types of CP violation

- **Observable CP violation in weak interaction can be classified into three different types**
  1. *Direct* CPV in **decays**
     - Observed in K, B, and D mesons
  2. *Indirect* CPV in **mixing**
     - Observed in K^0 oscillations
  3. CPV in decay/mixing **interference**
     - Observed in K^0, B^0 and B_s mesons

$$Pr(M \to f) \neq Pr(\overline{M} \to \overline{f})$$

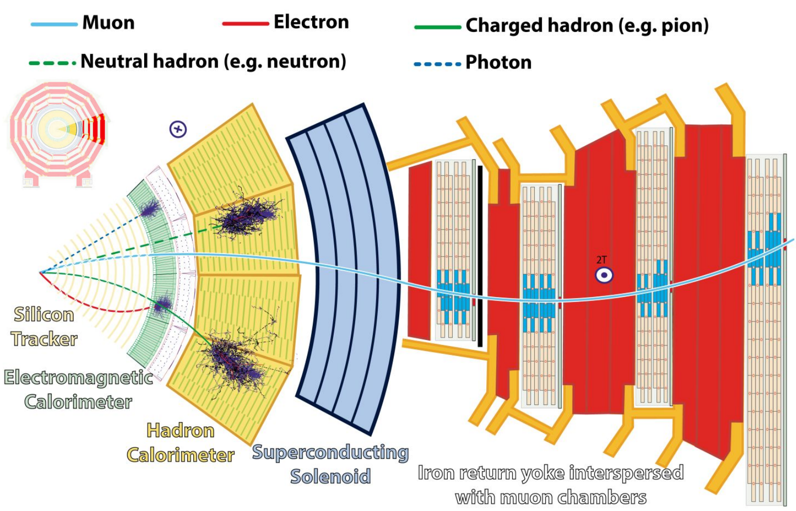
$$Pr(M^0 \to \overline{M}^0) \neq Pr(\overline{M}^0 \to M^0)$$

$$Pr(M^0_{(\rightsquigarrow \overline{M}^0)} \to f_{CP}) \neq Pr(\overline{M}^0_{(\rightsquigarrow M^0)} \to f_{CP})$$

- Defining A_f the M ➜ f amplitude, the CPV information can be coded in the **rephasing invariant complex parameter λ**

$$\boxed{\lambda \equiv \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_f}} \begin{cases} |\bar{A}_{\bar{f}}/A_f| \neq 1 & \to \text{direct CPV} \\ |q/p| \neq 1 & \to \text{indirect CPV} \\ |\lambda| = 1,\ \text{Im}(\lambda) \neq 0 & \to \text{interference CPV} \end{cases}$$

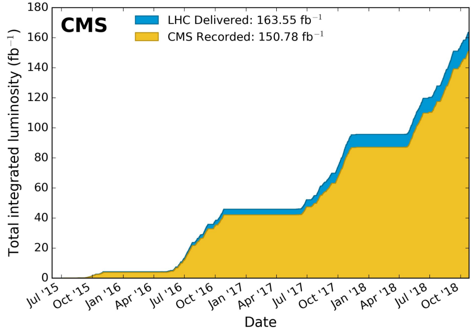
The CMS detector



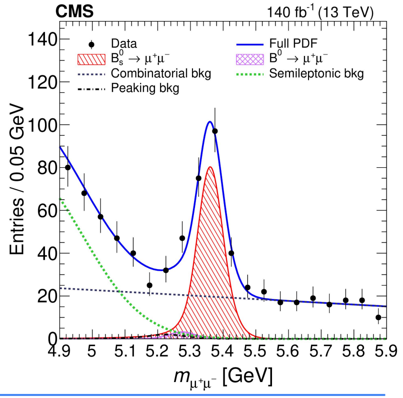
CMS is a general purpose detector able to perform a vast range of physics studies, including flavor physics

- \+ **Excellent tracking system** able to reconstruct vertices with high decay time resolution (e.g., $\boldsymbol{\sigma_t}$ **~ 65 fs** for $B_s \to J/\psi\ \phi$) up to $|\eta| < 2.5$
  - Complementary to LHCb ($2 < |\eta| < 5$)
- \+ **Enormous amount of data collected**
  - ~ $7.5 \cdot 10^{13}$ bb pairs produced at Point 5 during Run 2 (geometric acceptance not considered)
- – High pile up N_{PV} ~ 40 (in Run 2)
- – No reliable hadronic particle identification available

CMS luminosity in Run-2



$B_s \to \mu^+\mu^-$



Some **CMS flavor physics highlights** from recent years

- $B_s \to \mu^+\mu^-$ (world's most precise) [PLB842(2023)137955]
- $\eta \to \mu^+\mu^-\mu^+\mu^-$ observation [PRL131(2023)091903]
- f_s/f_u measurements [PRL131(2023)121901]
- Triple J/ψ production observation [Nat.Phys.19(2023)338]
- R(K) LFU test [BPH-22-005]
- R(J/ψ) LFU test [BPH-22-012]

Search for CP violation in $D^0 \rightarrow K_S K_S$

CMS PAS BPH-23-005

Dataset: 2018 B Parking (41 fb^{-1})

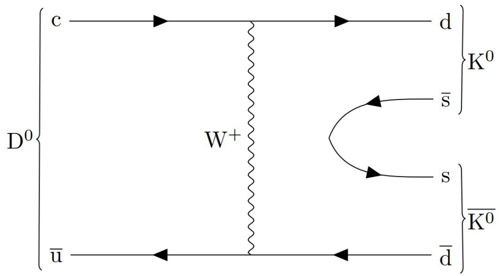
Motivations

- **CP violation in the up-quark sector is not studied as well as in the down-quark one**
  - Expected to be suppressed by the GIM mechanism and CKM element size
- **Observation of a significant CPV ➜ hints of BSM physics**
  - First observation of CPV in D mesons in 2019 by LHCb with $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays [PRL122(2019)211803]
- **This seminar presents a measurement of the direct CPV in $D^0 \to K_S K_S$ decays**

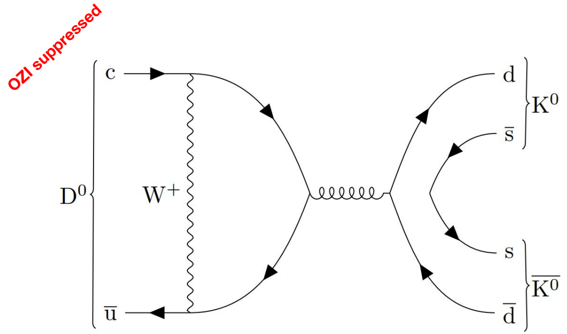
$$A_{CP} = \frac{\Gamma(D^0 \to K^0_S K^0_S) - \Gamma(\overline{D}^0 \to K^0_S K^0_S)}{\Gamma(D^0 \to K^0_S K^0_S) + \Gamma(\overline{D}^0 \to K^0_S K^0_S)}$$

- From theory, CPV in $D^0 \to K_S K_S$ could be as large as *O(1%)* [PRD92(2015)054036]

W exchange diagram for D^0



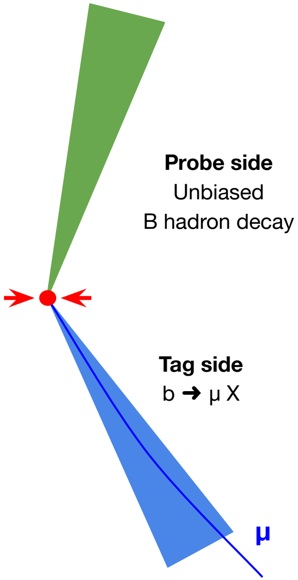
Penguin annihilation diagram for D^0



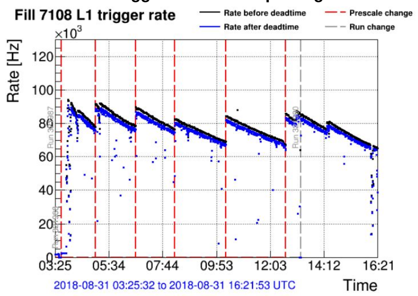
References
[CMS-EXO-23-007]
[CMS-BPH-22-005]

The CMS B parking dataset

- **Designed to allow CMS to perform B physics measurements on difficult/impossible to trigger final states** (e.g. fully hadronic final states)
- Achieved with a set of **single muon triggers** (tags) with different thresholds in p_T and impact parameter
  - Luminosity decreases during a run ➜ less restrictive triggers enabled
    - Maximises the available trigger bandwidth
  - Events are *parked* for later reconstruction
  - Very high purity of ~80%
- No impact on the *standard* CMS physics programme
- **10 billion unbiased B hadron** decays collected in 2018 (L_{int} ~ 41 fb^{-1})



L1 trigger rates with B parking



Measurement strategy

- **Use D^0 from $D^{*+} \rightarrow D^0 \pi^+$ and $D^{*-} \rightarrow \overline{D}^0 \pi^-$**, so that the pion charge tags the D^0 flavor
- This introduces the D^{*+}/D^{*-} differences in the measurement

What we want

$$A_{CP} = A_{raw} - A_{prod} - A_{det}$$

What we measure

$$A_{raw} = \frac{N(D^0) - N(\overline{D}^0)}{N(D^0) + N(\overline{D}^0)}$$

$$A_{prod} = \frac{\sigma_{pp \to D^{*+}X} - \sigma_{pp \to D^{*-}X}}{\sigma_{pp \to D^{*+}X} + \sigma_{pp \to D^{*-}X}}$$??

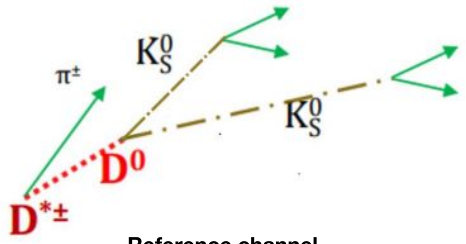
$$A_{det} \approx \frac{\epsilon_{\pi^+} - \epsilon_{\pi^-}}{\epsilon_{\pi^+} + \epsilon_{\pi^-}}$$??

- **Strategy: measure $\Delta A_{CP} = A_{CP}(D^0 \rightarrow K_S K_S) - A_{CP}(D^0 \rightarrow K_S \pi^+ \pi^-)$**
  - Reference channel is very similar in kinematics and topology ➜ A_{prod} and A_{det} cancel out
  - CPV in $D^0 \rightarrow K_s \pi^+\pi^-$ already measured consistent with zero [PRD86(2012)032007]

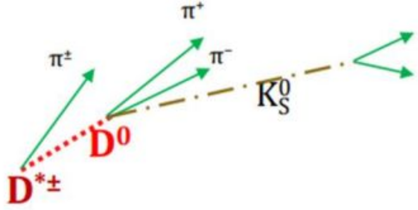
$$\Delta A_{CP} = A_{raw}(D^0 \to K_S K_S) - A_{raw}(D^0 \to K_S \pi^+ \pi^-)$$

Event selection

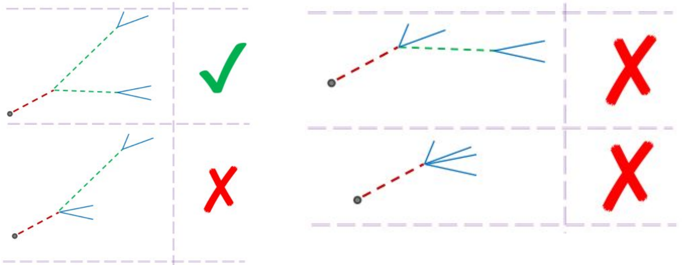
Signal channel



Reference channel



- **First, K_S ➜ $\pi^+\pi^-$** are reconstructed fitting the π tracks to a common vertex
  - $|m(\pi^+\pi^-) - m(K_S^{w.a.})| < 20$ MeV, $p_T(K_S) > 2.2(1.0)$ GeV
- In the **signal channel**, two K_S candidates are required and fitted to a common vertex to form D^0 ➜ $K_S K_S$ candidates
  - 1.7 GeV < $m(K_S K_S)$ < 2.0 GeV
  - K_S displacement in *xyz* from the D^0 vertex >9σ and >7σ
  - D^0 displacement in *xyz* (*xy*) from the PV >9σ (>2σ)
- In the **reference channel**, two track with $p_T > 0.6$ GeV are used to form the D^0 ➜ $K_S\pi^+\pi^-$ candidate
  - 1.823 < $m(K_S\pi^+\pi^-)$ < 1.908 GeV
- **Finally**, an additional track with $-1.2 < |\eta| < 1.2$ and $p_T > 0.36$ GeV is added to form D^{*+} ➜ $D^0\pi^+$ candidates
  - $m(D^0\pi^+) = m(D^0\pi^+) - m(D^0) + m_{PDG}(D^0)$



- **Background suppression**: several fits corresponding to incorrect topologies are performed and vertex probabilities requirements are imposed

A_{CP} extraction

To extract the CP asymmetry a **2D maximum-likelihood fit is performed on the invariant mass of the D^{*+} and D^0 invariant mass**

- Simultaneous on the D^{*+} and D^{*-} samples with only the yields left to float
- **Main fit components** (signal channel):
  - D^0 x D^{*+}, the signal component
  - D^0 x *bkg*, real D^0 and background pion
  - *bkg* x *bkg*, background in both dimensions
- **Signal pdf**: 2 x Johnson's SU
- **Background pdf**: polynomials, exponentials, threshold function
- **Yields:**

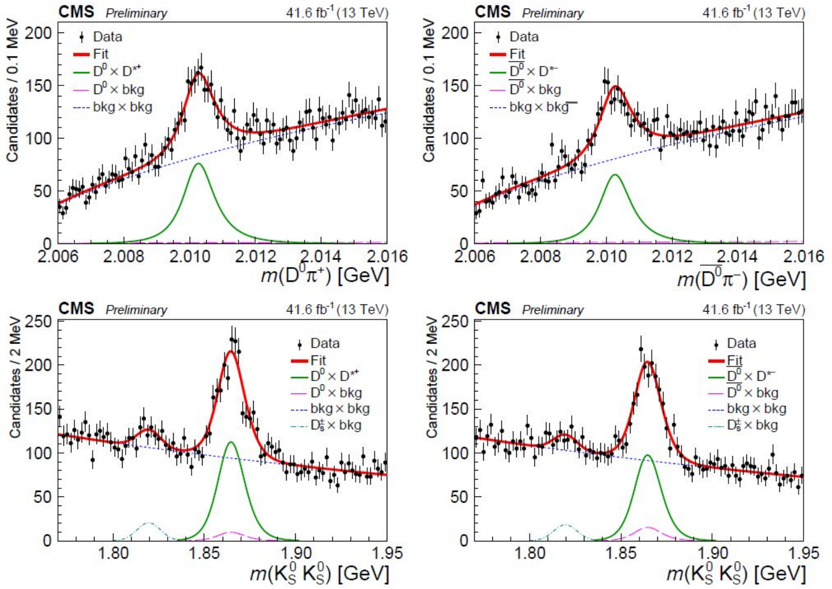
Reference

Pion charge	N
π^+	$944\,800 \pm 3\,500$
π^-	$930\,150 \pm 3\,400$

Signal

Pion charge	N
π^+	1095 ± 46
π^-	951 ± 44

$D^0 \to K_S K_S$ $\quad$ $\overline{D}^0 \to K_S K_S$



Systematic uncertainties

Source	Uncertainty, %
$m(D\pi^{\pm})$ signal model	0.10
$m(D\pi^{\pm})$ background model	0.02
$m(K^0_S K^0_S)$ signal model	0.04
$m(K^0_S K^0_S)$ background model	0.02
$m(K^0_S K^0_S)$ fit range	0.04
Reweighting	0.09
ΔA_{CP} in MC	0.13
Total	0.20

Results and outlook

- Putting everything together, ΔA_{CP} is measured

$$\Delta A_{CP} = 6.3 \pm 3.0\,(\text{stat}) \pm 0.2\,(\text{syst})\,\%$$

- **Using the world-average value of $A_{CP}(K_S\,\pi^+\pi^-) = (-0.1 \pm 0.8)\%$, $A_{CP}(K_SK_S)$ is found to be**

$$A_{CP}(D^0 \to K^0_S K^0_S) = 6.2 \pm 3.0\,(\text{stat}) \pm 0.2\,(\text{syst}) \pm 0.8(A_{CP}(K^0_S\,\pi^+\pi^-))\,\%$$

- Consistent with no CP violation at 2σ, with LHCb [PRD104(2021)L031102] [(-3.1 ± 1.3)%] at 2.7σ and Belle [PRL119(2017)171801] [(0.0 ± 1.5)%] at 1.8σ
- **This is the first CMS study of CP violation in the charm sector, paving the way for future measurements**
  - More data
  - Refined techniques
  - Different channels

Measurement of the $B_s \rightarrow J/\psi K_S$ effective lifetime

CMS PAS BPH-22-001

Dataset: 2016-18 (140 fb^{-1})

Motivations

- B_s mesons are produced in flavor eigenstates, but propagate as mass ones, which, if no **CPV** in the mixing, coincide with CP eigenstates

$$B_s^H \to \text{CP odd} \qquad B_s^L \to \text{CP even}$$

- These can have different lifetimes (as for the B_s), allowing the probe of the **mass eigenstate rate asymmetry $A_{\Delta\Gamma}$, directly related to the CPV observable λ**

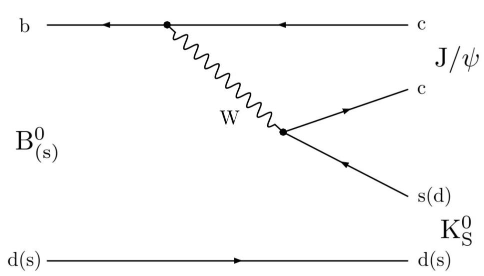
$$A_{\Delta\Gamma} = \frac{R_H - R_L}{R_H + R_L} = \frac{-2\,\mathcal{R}(\lambda)}{1 + |\lambda|^2}$$

  - R_H and R_L are related to the untagged decay rate as

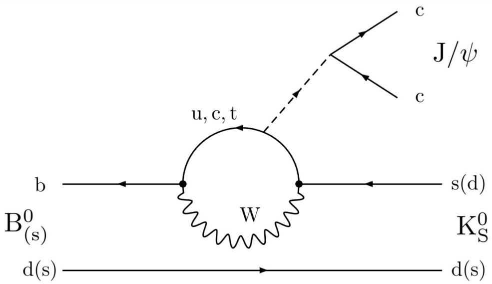
$$\Gamma(B_s \to f) + \Gamma(\overline{B}_s \to f) = R_H e^{-\Gamma_H t} + R_L e^{-\Gamma_L t}$$

- **This seminar presents a measure of the B_s effective lifetime τ in the CP-odd final state J/ψ K_S performed with the CMS Run 2 data set**
- This process is related to B^0 ➜ J/ψK_S via U-spin flavor symmetry
  - $A_{\Delta\Gamma}$ can be used to determine penguin contributions to the measurement of sin(2β)
  - The CKM angle γ can also be probed in B_S ➜ J/ψK_S

Tree-level diagram



Penguin loop diagram



The effective lifetime

- **The effective lifetime is defined as the expected value of the untagged decay rate**

$$\tau(J/\psi K_S) \equiv \frac{\int_0^\infty t(\Gamma_{B_s\to J/\psi K_S} + \Gamma_{\overline{B}_s\to J/\psi K_S})dt}{\int_0^\infty (\Gamma_{B_s\to J/\psi K_S} + \Gamma_{\overline{B}_s\to J/\psi K_S})dt} = \frac{\tau_{B_s}}{1-y_s^2}\left(\frac{1+2A_{\Delta\Gamma}y_s + y_s^2}{1+A_{\Delta\Gamma}y_s}\right)$$

- **Using the latest measurements and assuming the SM** ($A_{\Delta\Gamma}$ = 0.94 ± 0.07, τ_{Bs} = 1.520 ± 0.005 ps, $\Delta\Gamma$ = 0.084 ± 0.005 ps^{-1})

$$\tau(J/\psi K_S)\big|_{SM} = 1.62 \pm 0.02 \text{ ps}$$

  - Available measurement from LHCb: $\tau(J/\psi K_S)$ = 1.75 ± 0.14 ps [Nucl.Phys.B(2013)873]
- In this analysis the decay time is measured in the transverse plane as

$$t = \frac{L_{xy} \cdot M_{B_s}}{p_T}$$

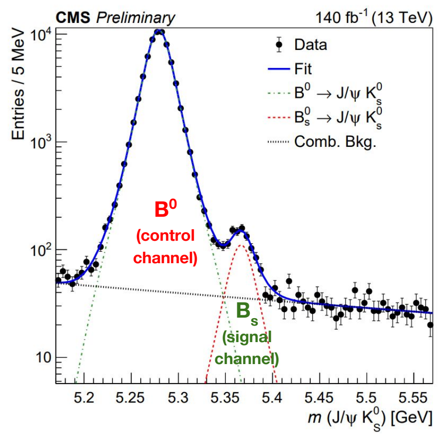
Event selection and efficiency

- **Trigger**: $J/\psi \rightarrow \mu^+\mu^-$ candidate with p_T > 20 (25) GeV for 2016 (2017-18)
- **Offline $K_S \rightarrow \pi^+\pi^-$ selection**
  - Displaced by >15σ from the beamspot and >5σ from the B_s vertex
  - Invariant mass within 70 MeV from world-average value
- **Background sources**
  - $\Lambda \rightarrow p\pi^-$: suppressed with constraints on the decay kinematics
  - $B^0 \rightarrow J/\psi\, K_s$: irreducible, treated as a control channel
  - $B^0 \rightarrow J/\psi\, K^{*0}$: negligible
  - Combinatorial: suppressed with dedicated BDT selection
- **Time efficiency**
  - Measured in simulations for B_s and B^0 (control channel)

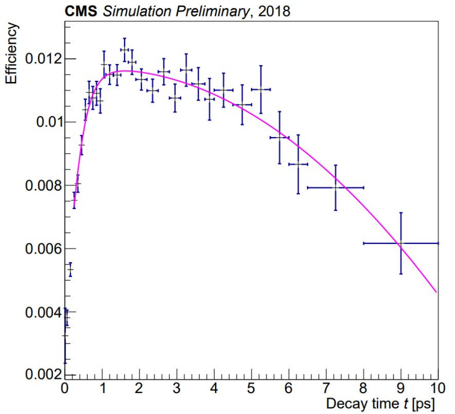
$$\epsilon(t) = \frac{t_{reco}}{t_{gen} \otimes \delta(t)}$$

  - Modeled with a combination of polynomials and logistic functions

Invariant mass distribution



B_s time efficiency (2018)



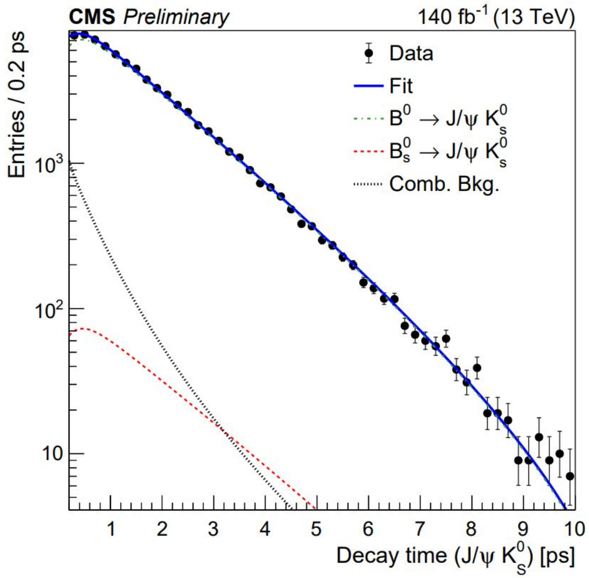
Fit and results

- **The effective lifetime is measured with a 2D UML fit to the invariant mass and proper decay time**
  - The decay time uncertainty is used as a conditional parameter
  - Both the effective lifetimes of the signal B_s and control channel B^0 are fitted
  - The control channel is used to validate most of the measurement components
- **Results** (using 727 ± 35 B_s signal candidates)

$$\tau(J/\psi\, K_S)^{eff} = 1.59 \pm 0.07\,(\text{stat}) \pm 0.03\,(\text{syst})\ \text{ps}$$

  - The control channel's effective lifetime is found to be in good agreement with the world-average value
- The measured $B_s \rightarrow J/\psi\, K_S$ effective lifetime is in agreement with the SM prediction and compatible with the previous LHCb results at 2.1σ
- **This is the most precise measurement of this quantity to date**

Proper decay time distribution



Systematic uncertainties

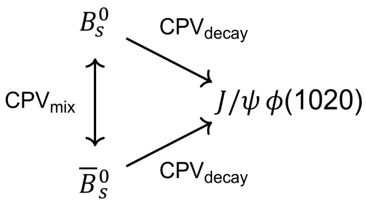
Source	Values (ps)
Deviation in control channel lifetime	0.002
Limited MC statistics	0.006
Efficiency modeling	0.002
Signal and background mass model	0.022
Background decay time model	0.014
Mass shape variation	0.007
Different fit strategy	0.006
Total	0.028

Measurement of the time-dependent CP violation in B_s mesons

CMS PAS BPH-23-004

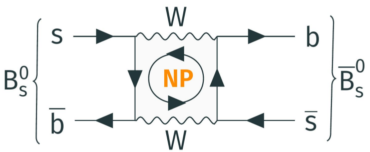
Dataset: 2017-18 (96 fb^{-1})

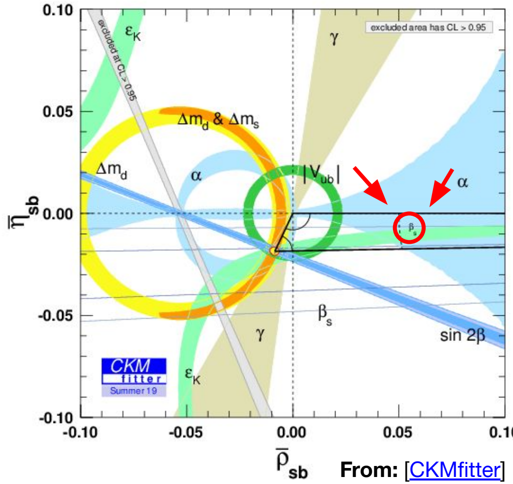
Motivations



$$\Gamma\left(B^0_{s\,(\rightsquigarrow \overline{B}^0_s)} \to f\right)(t) \overset{?}{\neq} \Gamma\left(\overline{B}^0_{s\,(\rightsquigarrow B^0_s)} \to f\right)(t)$$

- **B_s mesons decays allow us to study the time-dependent CP violation generated by the interference between direct decays and flavor mixing**
  - CPV in the interference is possible even if there is no CPV in decay and mixing
- **The weak phase ϕ_s is the main CPV observable**
  - Predicted by the SM to be $\boldsymbol{\phi_s \approx -2\beta_s}$ (β_s ➜ angle of the B_s unit. triangle)
    - Neglecting contributions from higher-order diagrams ($\Delta\phi_s^{loop} \approx 3\pm10$ mrad)
  - β_s determined by CKM global fits to be $\boldsymbol{-2\beta_s = -37 \pm 1}$ **mrad** [CKMfitter, UTfit]
- **New physics** can change the value of ϕ_s up to ~100% via new particles contributing to the flavor oscillations [RMP88(2016)045002]



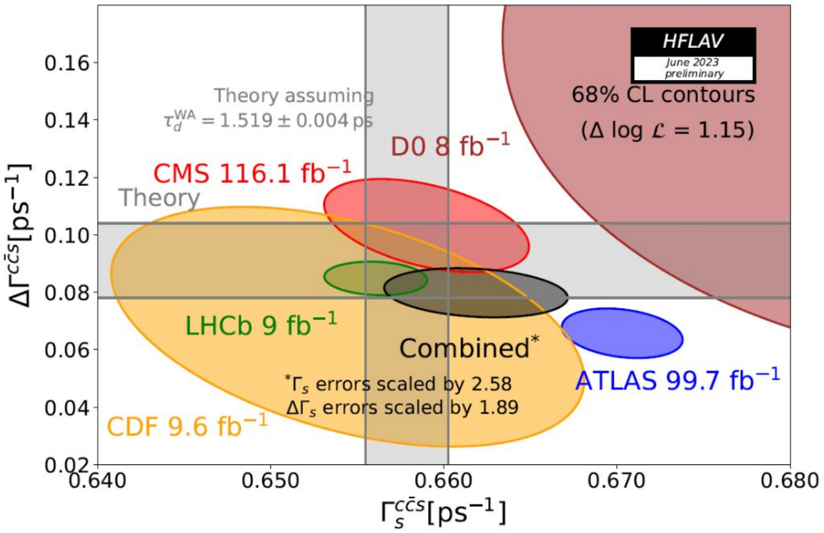


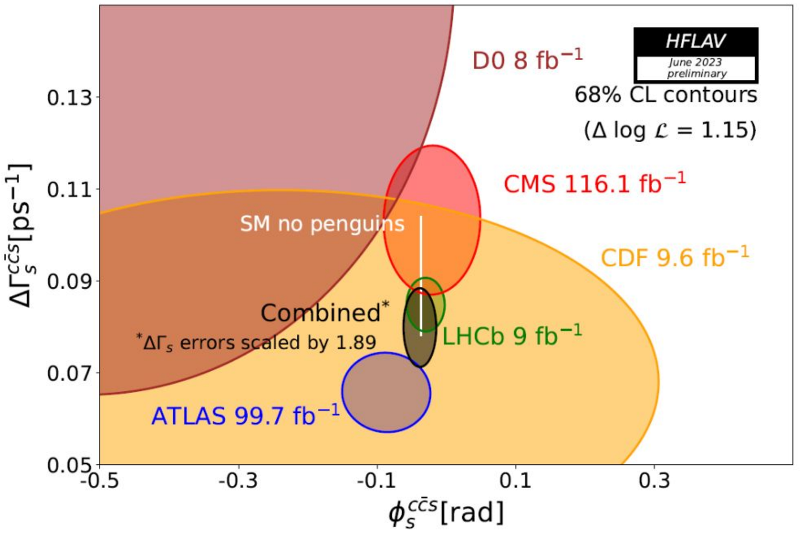
From: [CKMfitter]

- This seminar presents the latest CMS results with the *golden* channel $B_s \to J/\psi\,\phi(1020) \to \mu^+\mu^- K^+K^-$

A long history: flagship CPV analysis at LHC

- ϕ_s has been **first measured** by the **Tevatron** experiments D0 and CDF
- At LHC ϕ_s has been measured several times by ATLAS, LHCb, and CMS
  - LHCb has measured ϕ_s in several other channels, such as $B_s \to J/\psi\, \pi^+\pi^-$, $B_s \to J/\psi(e^+e^-)\, K^+K^-$, $B_s \to \psi\, K^+K^-$, $B_s \to D_s^+ D_s^-$, …
- Preliminary world-average (before this work): $\phi_s^{J/\psi KK} = -50 \pm 17$ mrad [JevticLiCERNSeminar(2023)]





From: [Jevtic and Li, CERN seminar (2023)]

A time-, flavor- and angular-dependent measurement

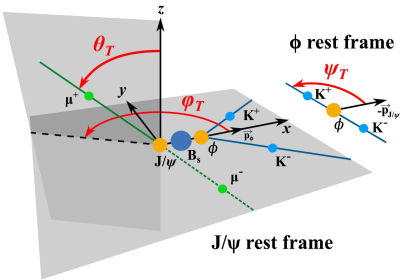
final-state CP eigenvalue

CP violation

flavor oscillations

$$a_{CP}(t) = \frac{-\eta_{fs}\sin(\phi_s)\sin(\Delta m_s t)}{\cosh(\frac{1}{2}\Delta\Gamma_s t) - \eta_{fs}\cos(\phi_s)\sinh(\frac{1}{2}\Delta\Gamma_s t)}$$

Transversity basis

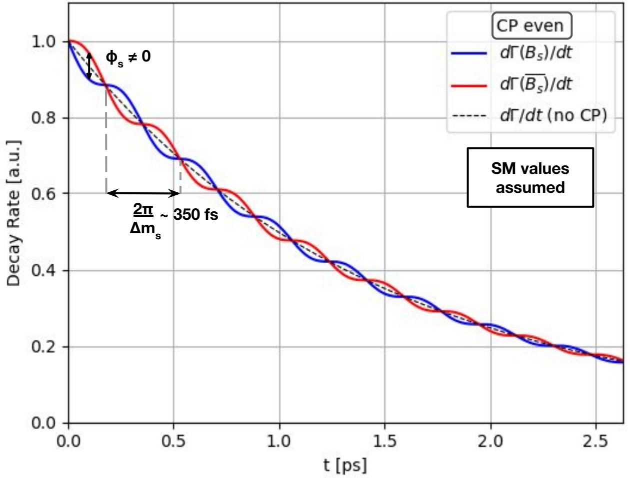


Core ingredients

- **Time-dependent angular analysis** to separate the CP eigenstates ("transversity basis" used)
- **Time-dependent flavor analysis** to resolve the B_s mixing oscillations (T ~ 350 fs)

$$\text{sensistivity} \propto \sqrt{\frac{\epsilon_{tag}\mathcal{D}^2_{tag}N_{sig}}{2}}\sqrt{\frac{N_{sig}}{N_{sig}+N_{bkg}}}\,e^{-\frac{\sigma_t^2\Delta m_s^2}{2}}$$

Decay rate for a CP-even final state



Decay rate model

Flavor tag decision (flips c_i and d_i signs)

Mistag probability

Decay time

Angular variables

$$\frac{d^4\Gamma(B_s)}{d\Theta dt} \propto \sum_{i=1}^{10} \mathcal{O}_i(t,\alpha)\, g_i(\Theta)$$

$$\mathcal{O}_i(t,\alpha) = N_i e^{-\Gamma_s t}\left[a_i \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + b_i \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + c_i \xi (1-2\omega)\cos(\Delta m_s t) + d_i \xi (1-2\omega)\sin(\Delta m_s t)\right]$$

Most sensitive terms for SM ϕ_s

i	$g_i(\theta_T, \psi_T, \varphi_T)$	N_i	a_i	b_i	c_i	d_i
1	$2\cos^2\psi_T(1-\sin^2\theta_T\cos^2\varphi_T)$	$\lvert A_0(0)\rvert^2$	1	D	C	$-S$
2	$\sin^2\psi_T(1-\sin^2\theta_T\sin^2\varphi_T)$	$\lvert A_\parallel(0)\rvert^2$	1	D	C	$-S$
3	$\sin^2\psi_T\sin^2\theta_T$	$\lvert A_\perp(0)\rvert^2$	1	$-D$	C	S
4	$-\sin^2\psi_T\sin 2\theta_T\sin\varphi_T$	$\lvert A_\parallel(0)\rvert\lvert A_\perp(0)\rvert$	$C\sin(\delta_\perp-\delta_\parallel)$	$S\cos(\delta_\perp-\delta_\parallel)$	$\sin(\delta_\perp-\delta_\parallel)$	$D\cos(\delta_\perp-\delta_\parallel)$
5	$\frac{1}{\sqrt{2}}\sin 2\psi_T\sin^2\theta_T\sin 2\varphi_T$	$\lvert A_0(0)\rvert\lvert A_\parallel(0)\rvert$	$\cos(\delta_\parallel-\delta_0)$	$D\cos(\delta_\parallel-\delta_0)$	$C\cos(\delta_\parallel-\delta_0)$	$-S\cos(\delta_\parallel-\delta_0)$
6	$\frac{1}{\sqrt{2}}\sin 2\psi_T\sin 2\theta_T\cos\varphi_T$	$\lvert A_0(0)\rvert\lvert A_\perp(0)\rvert$	$C\sin(\delta_\perp-\delta_0)$	$S\cos(\delta_\perp-\delta_0)$	$\sin(\delta_\perp-\delta_0)$	$D\cos(\delta_\perp-\delta_0)$
7	$\frac{2}{3}(1-\sin^2\theta_T\cos^2\varphi_T)$	$\lvert A_S(0)\rvert^2$	1	$-D$	C	S
8	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin^2\theta_T\sin 2\varphi_T$	$k_{SP}\lvert A_S(0)\rvert\lvert A_\parallel(0)\rvert$	$C\cos(\delta_\parallel-\delta_S)$	$S\sin(\delta_\parallel-\delta_S)$	$\cos(\delta_\parallel-\delta_S)$	$D\sin(\delta_\parallel-\delta_S)$
9	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin 2\theta_T\cos\varphi_T$	$k_{SP}\lvert A_S(0)\rvert\lvert A_\perp(0)\rvert$	$\sin(\delta_\perp-\delta_S)$	$-D\sin(\delta_\perp-\delta_S)$	$C\sin(\delta_\perp-\delta_S)$	$S\sin(\delta_\perp-\delta_S)$
10	$\frac{4}{3}\sqrt{3}\cos\psi_T(1-\sin^2\theta_T\cos^2\varphi_T)$	$k_{SP}\lvert A_S(0)\rvert\lvert A_0(0)\rvert$	$C\cos(\delta_0-\delta_S)$	$S\sin(\delta_0-\delta_S)$	$\cos(\delta_0-\delta_S)$	$D\sin(\delta_0-\delta_S)$

$$C = \frac{1-|\lambda|^2}{1+|\lambda|^2} \qquad S = -\frac{2|\lambda|\sin\phi_s}{1+|\lambda|^2} \qquad D = -\frac{2|\lambda|\cos\phi_s}{1+|\lambda|^2}$$

Sensitive to direct CPV

Sensitive to $\phi_s \sim 0$

Sensitive to $\phi_s \sim \pi/2$

Conventions

- $|A_{//}|^2 = |A_0|^2 - |A_\perp|^2$
- $\delta_0 = 0$
- $\delta_{S\perp} = \delta_S - \delta_\perp$
- $\Delta\Gamma_s > 0$

Physics parameters

- ϕ_s, $|\lambda|$
- $\Delta\Gamma_s$, Γ_s, Δm_s
- $|A_0|^2$, $|A_\perp|^2$, $|A_S|^2$
- $\delta_{//}$, $\delta_\perp$, $\delta_{S\perp}$

S-P wave effective coupling $k_{SP} \approx 0.54$

- Introduced since $m(K^+K^-)$ is not fitted
- Evaluated from the S- and P-wave lineshape interference

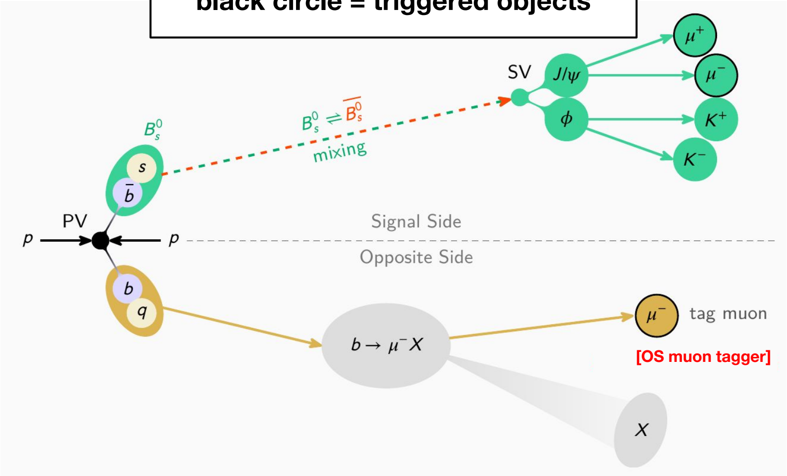
Trigger strategy

PLB816(2021)136188 (superseded)

black circle = triggered objects

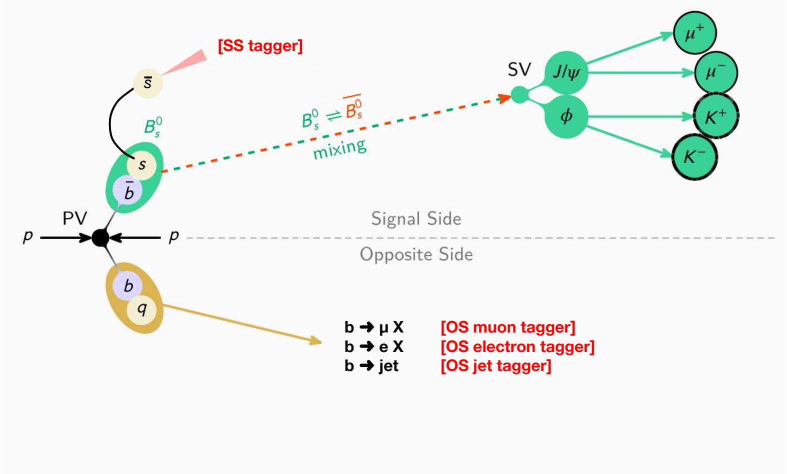
Muon-tagging trigger

- $J/\psi \rightarrow \mu^+\mu^-$ candidate plus an additional muon (for tagging)
- **≈50 000** signal candidates
- Used for time resolution modeling
- Tagging algorithms deployed: OS-muon
  - P_{tag} ~ **10%** (muon at trigger level enhance tagging efficiency)



Standard trigger

- Displaced $J/\psi \rightarrow \mu^+\mu^-$ candidate + $\phi(1020) \rightarrow K^+K^-$
- **≈450 000** signal candidates
- Tagging algorithms deployed: OS-muon, OS-electron, OS-jet, Same Side
  - P_{tag} ~ **5%**



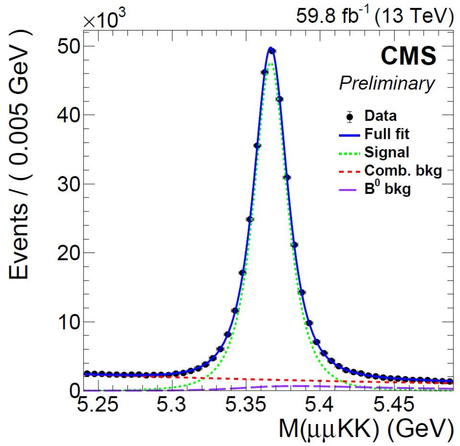
Dataset and selection

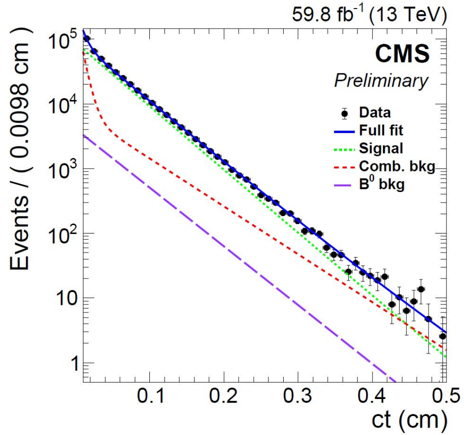
- **Dataset**: L_{int} = 96 fb^{-1} collected in 2017-2018
  - Why no 2016 data? Very different data set (old inner tracker detector with worse time resolution and different trigger menu)
- **Signal candidates: 491 270 ± 950**
- Notable selection requirements:

Variable	Requirement
ct (*muon-tagging* HLT)	> 60 μm
ct (*standard* HLT)	> 100 μm
ct/σ_{ct} (*standard* HLT)	> 3
$\lvert m(K^+K^-) - m_{\phi(1020)}\rvert$	< 10 MeV
$\lvert m(\mu^+\mu^-) - m_{J/\psi}\rvert$	< 150 MeV

- To avoid **overlaps**, events that pass both trigger category selections are placed only in the *muon-tagging* one
  - This depletes the *standard* trigger category of OS muons
- The PV of choice is the closest in 3D to the line that passes through the SV and parallel to the B_s momentum

Invariant mass and proper decay length distributions for the *standard* trigger (2018)





Decay time and its resolution

- The time dependence of the decay rate is parametrized with the **proper decay length** ct, measured in the transverse plane as

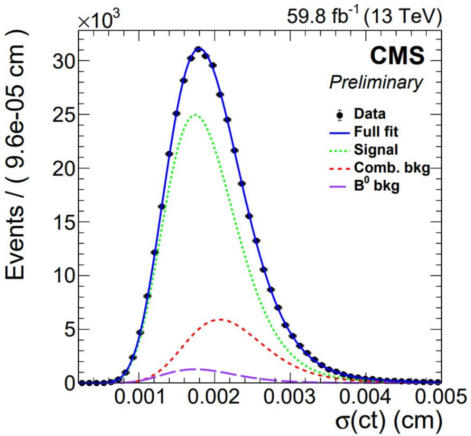
$$ct = c \cdot \frac{m_{Bs}^{w.a.} \cdot L_{xy}}{p_T} \quad \text{with} \quad L_{xy} \equiv ||\bar{r}_{xy}(SV) - \bar{r}_{xy}(PV)||$$

- Its **uncertainty** is obtained by fully propagating the uncertainties in L_{xy} and p_T
  - The uncertainty on L_{xy} dominates for most of the ct spectrum, with $\sigma(p_T)$ taking over at high values (ct ≳ 3 mm)
- **The ct uncertainty is calibrated in a prompt data sample** of $B_s \to J/\psi\ \phi$, obtained by removing the displacement requirement in the *muon-tagging* data sets
  - Modeled with two gaussians to obtain the effective dilution and resolution

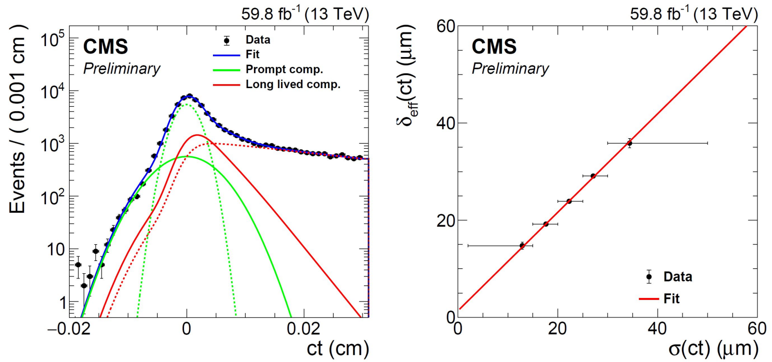
$$\delta_{\text{eff}} = \sqrt{\frac{-2 \ln \mathcal{D}}{\Delta m_s^2}} \quad \text{with} \quad \mathcal{D} = \sum_{i=1}^{2} f_i \exp\left(-\frac{\sigma_i \Delta m_s^2}{2}\right)$$

- **Excellent agreement** found, with corrections ~5%

Proper decay length uncertainty distribution for the *standard* trigger (2018)



Time resolution calibration for 2018 data



Acceptance and efficiency effects

- **The efficiency in selecting and reconstructing the B_s candidates is** **not** **independent of the decay time and angular observables**
  - To properly fit the decay rate model an efficiency parametrization is needed

Time efficiency

- Modeled in the $B^0 \rightarrow J/\psi\, K^{*0}$ data control channel with corrections from simulations
- Ultimately parametrized with Bernstein's polynomials

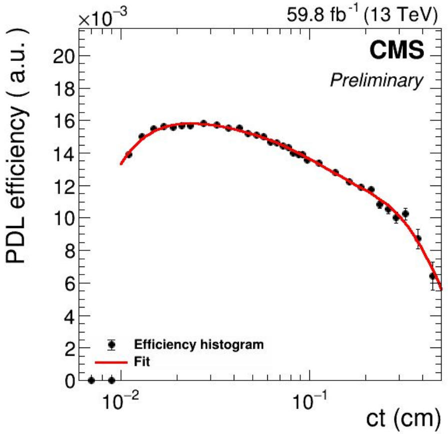
$$\varepsilon^{\text{data}}_{B^0}(ct) = \frac{N_{B^0}(ct)}{e^{-\Gamma^{\text{w.a.}}_{d}} \otimes P_{B^0}(\sigma_{ct})}$$

$$\varepsilon^{\text{data}}_{B_s}(ct) = \varepsilon^{\text{data}}_{B^0}(ct) \cdot \frac{\varepsilon^{\text{MC}}_{B_s}(ct)}{\varepsilon^{\text{MC}}_{B^0}(ct)}$$

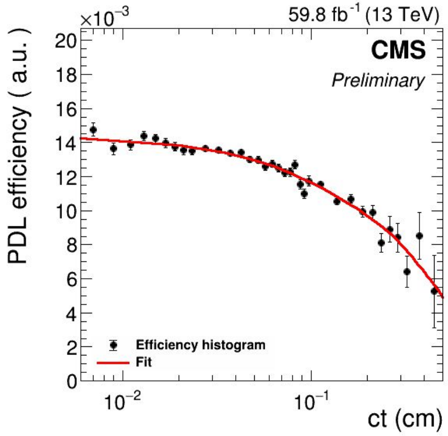
Angular efficiency

- Estimated with KDE distributions in simulated events
- The simulated data samples are corrected to match the data
  - An iterative procedure is used to simultaneously correct the kinematics of the final state particles and the differences in the physics parameters set in the MC with respect to what measured in the data

ct efficiency for the *standard* trigger category (2018)



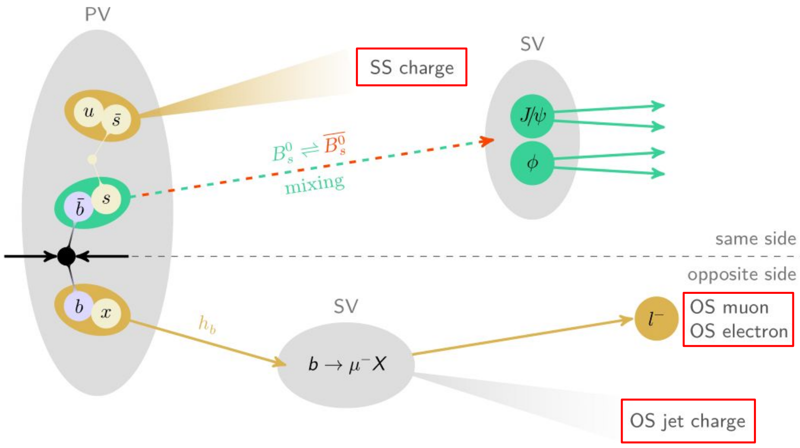
ct efficiency for the *muon-tagging* trigger category (2018)



Flavor tagging overview

- A **cutting-edge** **flavor tagging framework** has been engineered to extract the best possible results from data
- **Four DNN-based algorithms are used**, divided into two main categories
  - **Same side (SS):** exploits the B_s fragmentation
    1. **SS tagger**: leverages charge asymmetries in the B_s fragmentation
  - **Opposite side (OS)**: exploits decay products of the other B hadron in the event
    2. **OS muon**: leverages $b \rightarrow \mu^- X$ decays
    3. **OS electron**: leverages $b \rightarrow e^- X$ decays
    4. **OS jet**: capitalizes on charge asymmetries in the OS *b*-jet
- Only the OS-muon tagger is applied in the *muon-tagging* trigger category
  - The OS-electron, OS-jet and SS are applied only to the *standard* trigger category

Schematic representation of a generic event



Useful definitions

$$\xi_{tag} = \begin{cases} +1 & \text{for } B_s \\ -1 & \text{for } \overline{B}_s \\ 0 & \text{if no tagging decision is made} \end{cases}$$

$$\epsilon_{tag} = \frac{N_{tag}}{N_{tot}}, \quad \omega_{tag} = \frac{N_{mistag}}{N_{tag}}, \quad \mathcal{D}_{tag} = 1 - 2\omega_{tag}, \quad P_{tag} = \epsilon_{tag}\mathcal{D}_{tag}^2$$

Flavor, neural networks, and probabilities

- The **tagging inference** **logic** differs between algorithms
  - **Lepton taggers** (OS muon, OS electron)
    - Lepton charge ➜ ξ_{tag}; DNN score ➜ ω_{tag} *(DNN trained for correct-tag vs mistag)*

$$\text{OS } \ell^- \rightarrow \text{OS } b \xrightarrow{tag} \text{signal } B_s$$
$$\text{OS } \ell^+ \rightarrow \text{OS } \overline{b} \xrightarrow{tag} \text{signal } \overline{B}_s$$

$$\omega_{tag} = 1 - s_{DNN}$$

DNN score

  - **Charge-based taggers** (OS jet, SS)
    - DNN score ➜ Prob(B_s) ➜ ξ_{tag}, ω_{tag} *(DNN trained for B_s vs $\overline{B}_s$)*

$$s_{DNN} > 0.5 + \epsilon \xrightarrow{tag} \text{signal } B_s \quad \text{with} \quad \omega_{tag} = 1 - s_{DNN}$$
$$s_{DNN} < 0.5 - \epsilon \xrightarrow{tag} \text{signal } \overline{B}_s \quad \text{with} \quad \omega_{tag} = s_{DNN}$$

    - ϵ is used to remove events with ω_{tag} ~ 50%
- **The algorithms are optimized and trained in simulated events and calibrated in data with self-tagging $B^+ \rightarrow J/\psi K^+$ decays**
  - The calibration is performed by comparing ω_{tag} predicted by the DNN and the one measured in data

Calibration strategy (and other tricks)

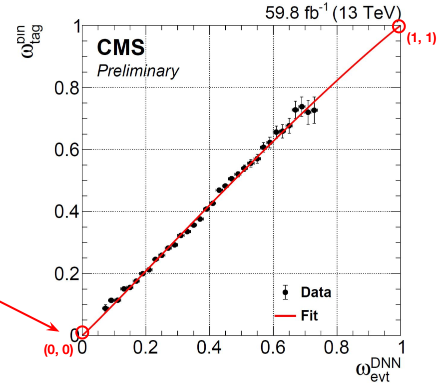
- **A multi-pronged strategy has been devised to improve the ω_{tag} estimation and suppress systematic effects**
  1. All models are constructed from the start as *probability estimators*, i.e. score~ω_{tag}
     - Loss function: *cross-entropy*, which is the likelihood for the probability P(true class | score)
     - Output layer: *Sigmoid* function, which normalizes the output to a probability distribution
  2. All DNNs are calibrated with the *Platt scaling*, which ensures that the calibrated score is still a probability
     - The Platt scaling is a linear calibration of the score before the last sigmoid layer
  3. In calibrating the charge-based taggers (which provide a probability for B_s vs $\overline{B}_s$):
     A. The output is *symmetrized* due to the initial LHC charge imbalance

$$s_{DNN}^{sym}(x) = \frac{s_{DNN}(x) + [1 - s_{DNN}(\overline{x})]}{2}$$

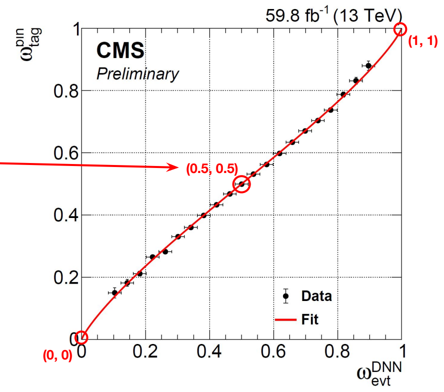
     B. The symmetry is explicitly forced in the calibration function by removing the constant term

This strategy cancels almost all the systematic effects associated with flavor tagging

OS-Muon calibration (*muon-tagging* trigger 2018)



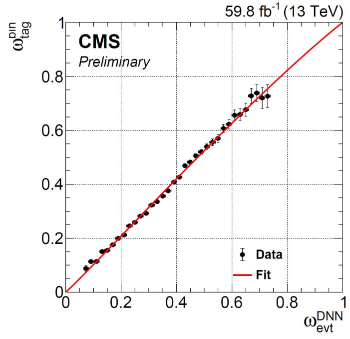
OS-jet calibration (2018)



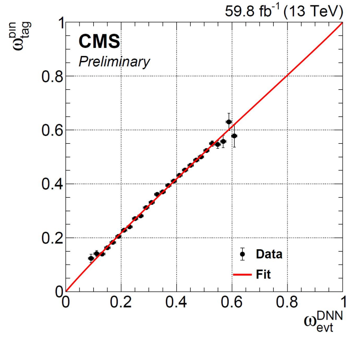
OS-lepton tagging

- **OS-lepton tagging techniques search for b ➜ ℓ^{-}X decays of the other B hadron in the event**
- The **charge** of the lepton is used as tagging feature and a fully connected DNN is used to estimate the mistag probability
- **Lepton selection**
  - Loose kinematic cuts
  - Separated from the signal B meson
  - MVA discriminator against fakes
  - OS-electrons are searched only if no OS-muon is found in the event (explicit orthogonality)
- **Mistag estimation**
  - Fully connected DNN with ReLU activation and dropout
  - Inputs: lepton kinematics and surrounding activity
- **Trained on simulated B_s ➜ J/ψ ϕ(1020) events and calibrated in B^+ ➜ J/ψ K^+ data**

OS-Muon calibration (*muon-tagging* trigger 2018)



OS-Electron calibration (2018)

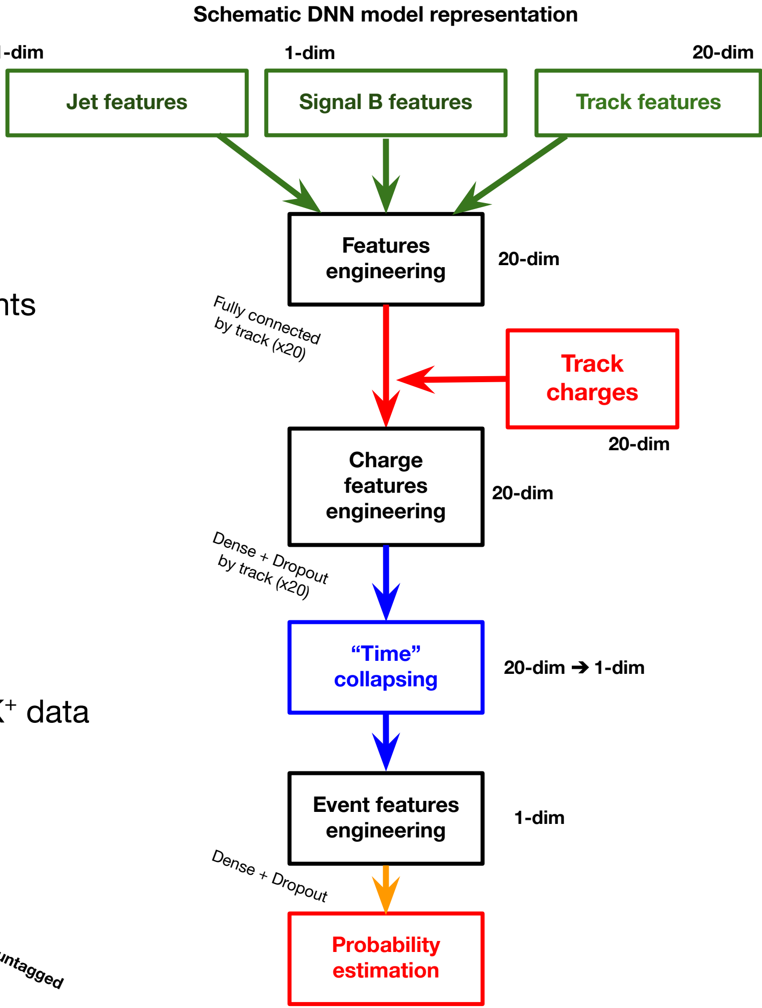


OS-jet tagging

- **The OS-jet algorithm exploits charge asymmetries in the jet structure and is based on a DNN called *DeepJetCharge***
  - Inputs: features from signal B meson, OS jet and its constituents
    - NB: The only flavor asymmetry is in the charges
  - Based on the *DeepSets* architecture [ref]
- **Jet selection**
  - No OS-lepton candidate
  - At least 2 tracks with $|IP_z| < 1$ cm
  - Separated from the signal B meson
  - jet b-tagging discriminator
- Additional nearby tracks are used due to the poor jet clustering performance in the kinematic region of interest ($p_T < 20$ GeV)
- Trained on simulated $B_s \to J/\psi\, \phi$ events and calibrated in $B^+ \to J/\psi\, K^+$ data
- **The trained network produces the probability of signal B meson containing a $\bar{b}$ quark (i.e. being a B_s)**
- The score is finally used to compute both ξ_{tag} and ω_{tag}

$$s_{DNN} > 0.52 \xrightarrow{tag} \text{signal } B_s \quad \text{with} \quad \omega_{tag} = 1 - s_{DNN}$$
$$s_{DNN} < 0.48 \xrightarrow{tag} \text{signal } \overline{B}_s \quad \text{with} \quad \omega_{tag} = s_{DNN}$$

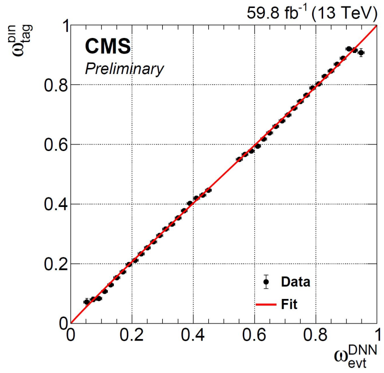
$\omega_{tag} > 0.48 \to$ untagged



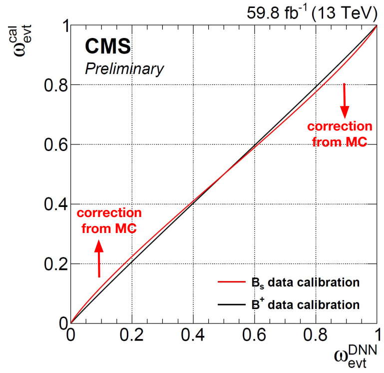
SS tagger

- **The SS tagger consists of a DNN (*DeepSSTagger*), derived from *DeepJetCharge,* able to probe the fragmentation products of a B meson and exploit tracks with high flavor correlation**
- *DeepSSTagger* uses the kinematic information from up to 20 tracks (ordered by $|IP_z|$) around the reconstructed B meson
- **Track selection**
  - $\Delta R(\text{trk}, B) < 0.8$, $|IP_z(PV)| < 0.4$ cm, $|IP_{xy}(PV)|/\sigma_{dxy} < 1$
  - Overlap with signal and OS is carefully avoided with geometrical cuts and vetos
- **Trained on an equal-weight mixture** of $B_s \to J/\psi\, \phi$ and $B^+ \to J/\psi\, K^+$ to make the model invariant for $B_s \leftrightarrow B^+$ for calibration purposes
  - Calibration directly in B_s was found to be not feasible in CMS
    - Tested: $B_s \to D_s^-\pi^+$ (not enough stat.) and $B_s^{**} \to B^{+(*)}K^-$ (too much uncer. from B^{0**} bkg)
  - The trained network produces the probability of signal B meson containing a negatively charged quark alongside the b quark (i.e., being a B_s or B^-)
- **Calibration**
  - The SS is calibrated $B^+ \to J/\psi\, K^+$ data, with residual differences ~10% corrected with simulations
  - Events with $\omega_{tag} > 0.46$ are removed before the calibration and assumed untagged

Same-side tagger calibration (B^+ data 2018)



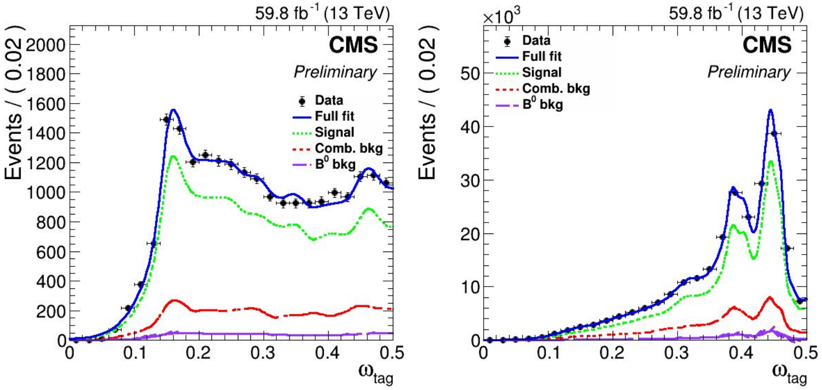
Comparison between Same-side tagger B^+ and B_s calibrations (2018)



Flavor tagging performance

- The SS and any one of the OS algorithms overlap in about 20% of the events
  - In these cases, the information is combined to improve the tagging inference
- **The combined flavor tagging framework achieves a tagging power of P_{tag} = 5.6%** when applied to the B_s data sample
  - Among the highest ever recorded at LHC
  - x3~4 improvement with respect to prev. CMS results
- **This is the first CMS implementation of the OS jet and same-side tagging techniques**
  - SS accounts for half of the performance
- Largest ever effective statistics $N_{Bs} \cdot P_{tag}$ (490k · 5.6% ≈ 27.5k) for a single ϕ_s measurement
- The flavor tagging framework is validated in the $B^0 \rightarrow J/\psi K^{*0}$ data control channel with flavor mixing measurements, both integrated and time-dependent

ω_{tag} distribution in the *muon-tagging* trigger category (left) and the *standard* one (right) for 2018 data



Flavor tagging performance (mutually exclusive categories)

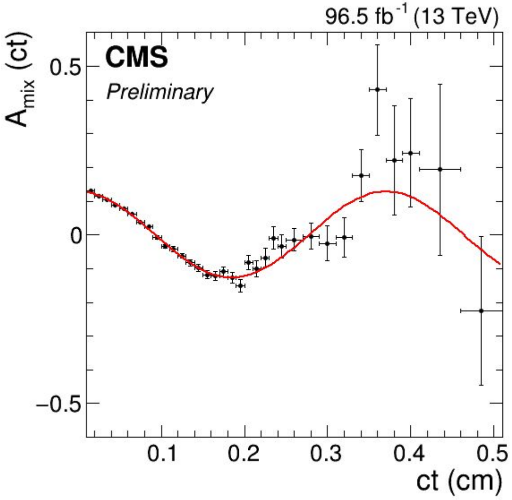
Category	ε_{tag} [%]	$\mathcal{D}^2_{eff}$	P_{tag} [%]
Only OS muon	6.07 ± 0.05	0.212	1.29 ± 0.07
Only OS electron	2.72 ± 0.02	0.079	0.214 ± 0.004
Only OS jet	5.16 ± 0.03	0.045	0.235 ± 0.003
Only SS	33.12 ± 0.07	0.080	2.64 ± 0.01
SS + OS muon	0.62 ± 0.01	0.202	0.125 ± 0.003
SS + OS electron	2.77 ± 0.02	0.150	0.416 ± 0.005
SS + OS jet	5.40 ± 0.03	0.124	0.671 ± 0.006
Total	55.9 ± 0.1	0.100	5.59 ± 0.02

Tagging validation with B^0 events

- **The flavor tagging framework is validated in the $B^0 \to J/\psi\, K^{*0}$ control channel** (~2M events)
- The time-dependent **mixing asymmetry** is measured to extract the flavor mixing oscillation frequency Δm_d with a precision of ~*1%* (comparable with BaBar and Belle)
  - Excellent agreement with world-averages is observed
    ➜ **No bias in mixing frequency measurements**
- Study performed also in each tagging category (see backup)
- The **time-integrated mixing** is also measured for each tagger and their dependency on the expected tagging dilution is compared
  - The dependency between the measured A_{mix} and the estimated D_{tag} is found to be well described by a linear relationship, indicating that all four techniques behave in the same predictable way

B^0 flavor mixing asymmetry

$$A_{mix}(ct) = \frac{N_{unmix}(ct) - N_{mix}(ct)}{N_{unmix}(ct) + N_{mix}(ct)}$$



Fit model

- The physics parameters are extracted with **unbinned multidimensional extended maximum-likelihood (UML) fit** performed simultaneously on **12 data sets** (2 trig. cat. x 2 years x 3 ξ_{tag} values)
  - *Physics parameters*: ϕ_s, $|\lambda|$, $\Delta\Gamma_s$, Γ_s, Δm_s, $|A_0|^2$, $|A_\perp|^2$, $|A_S|^2$, $\delta_{//}$, $\delta_\perp$, $\delta_{S\perp}$
  - *Observables*: m_{Bs}, ct, σ_{ct}, $\cos\theta_T$, $\cos\psi_T$, ϕ_T, ω_{tag}
- **Fit model**

$$P = f_{sig} P_{sig} + f_{bkg} P_{bkg} + f_{bkg\,B^0} P_{bkg\,B^0}$$

Angular eff

$B_s \to J/\psi\,\phi$ decay rate

Observables pdfs

SIGNAL
$$P_{sig} = \varepsilon(\Theta)\,[\tilde{f}(\Theta, ct \mid \alpha, \xi_{tag}, \omega_{tag}) \otimes G(ct, \sigma_{ct})]\, P_{sig}(m_{B_s})\, P_{sig}(\sigma_{ct})\, P_{sig}(\omega_{tag})$$

COMBINATORIAL BKG
$$P_{bkg} = [P_{bkg}(ct) \otimes G(ct, \sigma_{ct})]\, P_{bkg}(\Theta)\, P_{bkg}(m_{B_s})\, P_{bkg}(\sigma_{ct})\, P_{bkg}(\omega_{tag})$$

$B^0 \to J/\psi K^{0*}$ BKG
$$P_{bkg\,B^0} = [P_{bkg\,B^0}(ct) \otimes G(ct, \sigma_{ct})]\, P_{bkg\,B^0}(\Theta)\, P_{bkg\,B^0}(m_{B_s})\, P_{bkg\,B^0}(\sigma_{ct})\, P_{bkg\,B^0}(\omega_{tag})$$

Bkg time pdf

Time resolution convolution

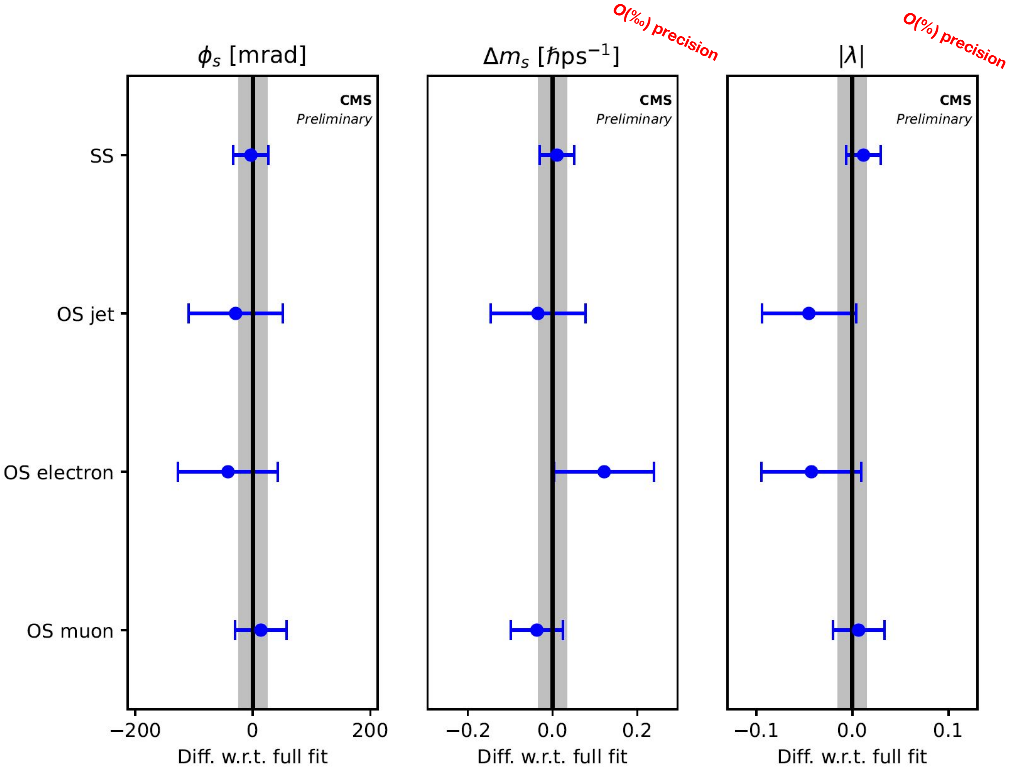
- The time efficiency is implemented as a *re-weighting* of the data events to drastically improve fit time
- The statistical uncertainties and fit bias are estimated with **1300 *bootstrap* distributions**
- The yield for the **$B^0 \to J/\psi K^{*0}$ is estimated directly in data** with a 2D fit to the B_s invariant mass and its B^0 reflection
- The background from **$\Lambda_b \to J/\psi K^- p^+$ is found to be negligible** and is treated as a systematic uncertainty

Systematic uncertainty overview

	ϕ_s [mrad]	$\Delta\Gamma_s$ [ps^{-1}]	Γ_s [ps^{-1}]	Δm_s [$\hbar$ps^{-1}]	$\lvert\lambda\rvert$	$\lvert A_0\rvert^2$	$\lvert A_\perp\rvert^2$	$\lvert A_S\rvert^2$	$\delta_\parallel$ [rad]	$\delta_\perp$ [rad]	$\delta_{S\perp}$ [rad]
Statistical uncertainty	23	0.0043	0.0015	0.035	0.014	0.0016	0.0021	0.0033	0.074	0.089	0.15
Model bias	4	0.0011	0.0002	0.004	0.006	0.0012	0.0022	0.0006	0.015	0.017	0.03
Flavor tagging	4	$<10^{-4}$	0.0005	0.007	0.002	$<10^{-4}$	$<10^{-4}$	0.0006	0.012	0.016	0.03
Angular efficiency	4	0.0002	$<10^{-4}$	0.015	0.011	0.0042	0.0019	0.0001	0.017	0.044	0.02
Time efficiency	<1	0.0014	0.0026	$<10^{-3}$	$<10^{-3}$	0.0004	0.0005	$<10^{-4}$	0.001	0.002	$<10^{-2}$
Time resolution	<1	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$	$<10^{-3}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$	0.001	$<10^{-3}$
Model assumptions	—	0.0005	0.0006	—	—	—	—	—	—	—	—
B^0 background	<1	0.0002	0.0003	$<10^{-3}$	$<10^{-3}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$	$<10^{-3}$	$<10^{-2}$
Λ_b^0 background	—	—	0.0004	—	—	0.0004	0.0003	—	—	—	—
S-P wave interference	<1	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$	$<10^{-3}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$	$<10^{-3}$	$<10^{-2}$
$P(\sigma_{ct})$ uncertainty	<1	0.0002	0.0003	$<10^{-3}$	$<10^{-3}$	0.0001	0.0001	$<10^{-4}$	$<10^{-3}$	$<10^{-3}$	$<10^{-2}$
Total systematic uncertainty	7	0.0019	0.0028	0.017	0.012	0.0044	0.0030	0.0009	0.025	0.050	0.05

- **Model bias, flavor tagging, and angular efficiency are found to be the leading systematic sources for ϕ_s**
- The measurement is still heavily statistically limited for ϕ_s

Validation: fit with individual tagging techniques



- **To check the consistency and stability of the tagging framework, the fit to data is repeated with only one tagging algorithm deployed at a time**
  - The grey area represents the result and statistical uncertainty of the full fit
  - Only flavor-sensitive parameters are presented
- **Excellent agreement between the various tagging techniques**

Results

Fit results

Parameter	Fit value	Stat. uncer.	Syst. uncer.
ϕ_s [mrad]	-73	± 23	± 7
$\Delta\Gamma_s$ [ps^{-1}]	0.0761	± 0.0043	± 0.0019
Γ_s [ps^{-1}]	0.6613	± 0.0015	± 0.0028
Δm_s [$\hbar$ps^{-1}]	17.757	± 0.035	± 0.017
$\lvert\lambda\rvert$	1.011	± 0.014	± 0.012
$\lvert A_0\rvert^2$	0.5300	± 0.0016	± 0.0044
$\lvert A_\perp\rvert^2$	0.2409	± 0.0021	± 0.0030
$\lvert A_S\rvert^2$	0.0067	± 0.0033	± 0.0009
$\delta_\parallel$	3.145	± 0.074	± 0.025
$\delta_\perp$	2.931	± 0.089	± 0.050
$\delta_{S\perp}$	0.48	± 0.15	± 0.05

- **ϕ_s and $\Delta\Gamma_s$ are found in agreement with the SM**

$$\phi_s^{SM} \simeq -37 \pm 1 \text{ mrad} \qquad \Delta\Gamma_s^{SM} = 0.091 \pm 0.013 \text{ ps}^{-1}$$

- **Γ_s and Δm_s are consistent with the latest world averages**

$$\Gamma_s^{WA} = 0.6573 \pm 0.0023 \text{ ps}^{-1} \qquad \Delta m_s^{WA} = 17.765 \pm 0.006\ \hbar\text{ps}^{-1}$$

- **$|\lambda|$ is consistent with no direct CPV ($|\lambda| = 1$)**

- This measurement utilizes the **largest ever** effective statistics $N_{Bs} \cdot P_{tag}$ for a single ϕ_s measurement
    - The precision on ϕ_s is comparable with the world's most precise single measurement by LHCb (ϕ_s = -39 ± 22 (stat) ± 6 (syst) mrad) [PRL132(2024)051802]
    - This is the most precise single measurement of $\Delta\Gamma_s$ to date in this channel

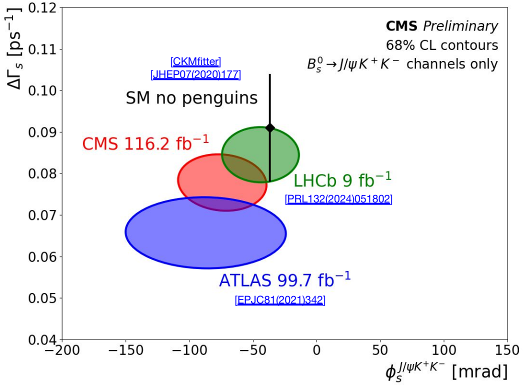
Combination with 8 TeV results

- These results supersede PLB816(2021)136188 and are further combined with those obtained CMS at 8 TeV [PLB757(2016)97], yielding

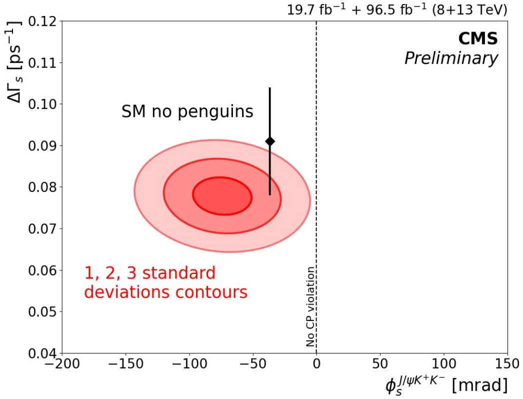
$$\phi_s = -74 \pm 23 \text{ [mrad]}$$
$$\Delta\Gamma_s = 0.0780 \pm 0.0045 \text{ [ps}^{-1}\text{]}$$

- Due to the high difference in statistical power between the two results the sensitivity gain is small
- **The combined value for the weak phase ϕ_s is consistent with the SM prediction, the latest world average, and with zero (no CPV) at 3.2 s.d.**
  - **This is the first evidence of CPV in $B_s \to J/\psi K^+K^-$ decays**
- These results helps to further constrain possible BSM effects in the B_s system

Comparison with other LHC experiments



1, 2, 3 standard deviations contours



Outlook

Summary and outlook

- This seminar presented three recent CMS results on the physics of CP violation
  - **CP violation in $D^0 \rightarrow K_S K_S$**
    - First CMS results on CP violation in the charm sector
  - **Effective lifetime measurement in the CP-odd decay $B_s \rightarrow J/\psi\, K_S$**
    - Most precise determination of $\tau_{eff}(B_s \rightarrow J/\psi\, K_S)$
  - **Measurement of the time-dependent CP violation in $B_s \rightarrow J/\psi\, \phi$**
    - First evidence of CP violation in $B_s \rightarrow J/\psi\, K^+K^-$
- CMS recent contributions in flavor physics prove that it can be one of the leading actors in several key areas of study, such as rare decays and CP violation
- Advancements in trigger strategies and flavor tagging techniques allows CMS to compete in measurements for which the detector was not designed
- **Run 3 will provide unique opportunities thanks of a revamped trigger strategy, which will lead to the collection of an unprecedented amount of data suitable for flavor physics studies**

Stay tuned in the future for other exciting CMS results!

Thanks for the attention

Backup
- CP violation in D^0 -

Selection

Table 1: Optimized selection criteria in the signal channel $D^0 \to K^0_S K^0_S$.

Variable	Requirement
p_T of tagging pion from $D^{*\pm} \to D\pi^\pm$	$> 0.35\,\text{GeV}$
η of tagging pion from $D^{*\pm} \to D\pi^\pm$	$-1.2 < \eta < 1.2$
$p_T(K^0_S)$	$> 2.2\,\text{GeV}$ and $> 1.0\,\text{GeV}$
$P_{vtx}(D\pi^\pm)$	$> 5\%$
$P_{vtx}(K^0_S K^0_S)$	$> 1\%$
$P_{vtx}(\pi^+\pi^-)$ for $K^0_S \to \pi^+\pi^-$	$> 1\%$
D^0 vertex displacement from the PV in xy	> 2 s.d.
D^0 vertex displacement from the PV in xyz	> 9 s.d.
K^0_S vertex displacement from the D^0 vertex in xyz	> 9 s.d. and > 7 s.d.
angle between D^0 momentum and displacement from PV in xyz	< 0.205 rad
angle between D^0 momentum and displacement from PV in xy	< 0.237 rad
angle between D^0 momentum and displacement from BX in xy	< 0.237 rad

Backup

- CP violation in B_s -

Penguin contributions

Assuming this is negligible

We measure this

$$\phi_s = \phi_s^{tree} + \Delta\phi_s^{penguin} + \Delta\phi_s^{NP}$$

$$\sin(2\beta) = \sin(\,2\beta^{tree} + \Delta\phi_d^{penguin} + \Delta\phi_d^{NP})$$

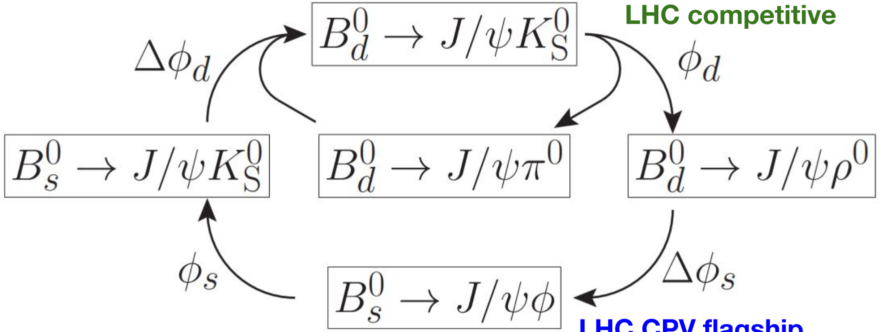
Trying to probe this

- Penguin pollutions are expected to be small for B_s, but they are not well constrained

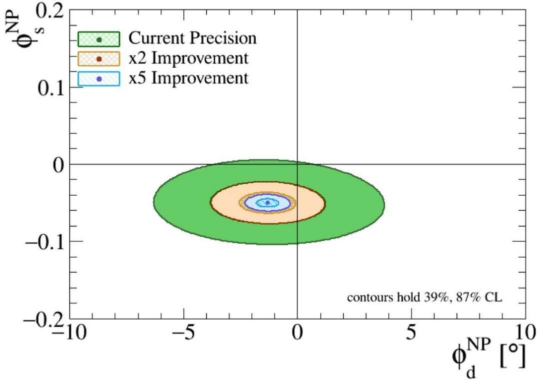
$$\Delta\phi_s^{penguin} \approx 3 \pm 10 \text{ mrad}$$

- Analysis of penguin and NP contributions is possible using Cabibbo-favored control channels

B-factories CPV flagship
LHC competitive



LHC CPV flagship
B-factories not competitive



Offline selection

Requirements common between the two HLTs

- $5.24 < m(\mu\mu KK) < 5.49$ GeV
- $p_T(B_s) > 9.5$ GeV
- Vertex probability > 2%
- $\sigma(ct) < 50\ \mu m$
- $|\eta(\mu)| < 2.4$
- $|\eta(K)| < 2.5$
- $|m(\mu\mu) - m(J/\psi^{PDG})| < 150$ MeV
- $|m(KK) - m(\phi(1020)^{PDG})| < 10$ MeV

Requirements specific to the *muon-tagging HLT*

- $p_T(\mu) > 3.5$ GeV
- $p_T(K) > 1.15$ GeV
- $ct > 60\ \mu m$

Requirements specific to the *standard* HLT

- *muon-tagging* trigger vetoed
- $p_T(\mu) > 4$ GeV
- $p_T(K) > 0.9$ GeV
- $p_T(\mu\mu) > 6.9$ GeV
- $ct > 100\ \mu m$, $ct/\sigma(ct) > 3$

- **Selection requirement** **optimized** with the a genetic algorithm to maximize $S/\sqrt{(S + B)}$
- **To avoid** **overlaps**, **the *muon-tagging* trigger is vetoed in the *standard* trigger category**
- The **PV** of choice is the closest in 3D to the line that passes through the SV and parallel to the B_s momentum

OS-lepton taggers selection

OS Muon

- **Requirements**
  - $p_T > 2$ GeV
  - $|\eta| < 2.4$
  - $|d_z(PV)| < 1$ cm
  - $\Delta R(B_s) > 0.4$
  - Discriminators vs fakes

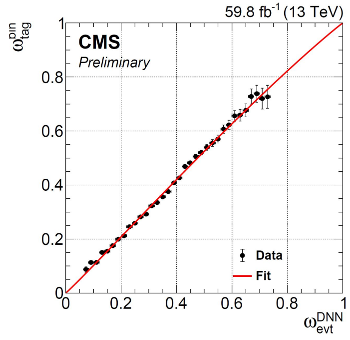
- Deployed in **both trigger categories**
- Dense DNN for ω_{tag} estimation
  - Inputs: kinematics, IP, surrounding activity

OS electron

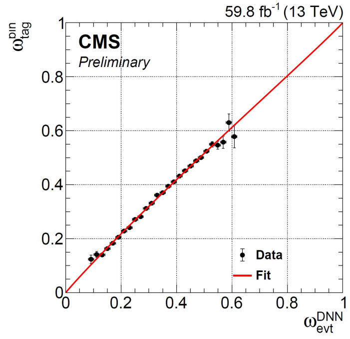
- **Requirements**
  - No OS muon selected in the event
  - $p_T > 2.5$ GeV
  - $|\eta| < 2.4$
  - $|d_z(PV)| < 0.2$ cm
  - $|d_{xy}(PV)| < 0.08$ cm
  - $\Delta R(B_s) > 0.4$
  - Discriminators vs fakes

- Deployed **only in the *standard* trigger** category
- Dense DNN for ω_{tag} estimation
  - Inputs: kinematics, IP, surrounding activity

OS-Muon calibration (*muon-tagging* HLT 2018)



OS-Electron calibration (2018)



Taggers combination

- **Overlap logic**

Overlap	OS muon	OS electron	OS jet	SS
OS muon		X	X	✓
OS electron	X		X	✓
OS jet	X	X		✓
SS	✓	✓	✓	

- **Tag decision combination**

$$\xi(\xi_1, \xi_2, \omega_1, \omega_2) = \begin{cases} \xi_1 & \text{if} \quad \omega_1 < \omega_2 \\ \xi_2 & \text{if} \quad \omega_2 < \omega_1 \end{cases}$$

- **Mistag combination**

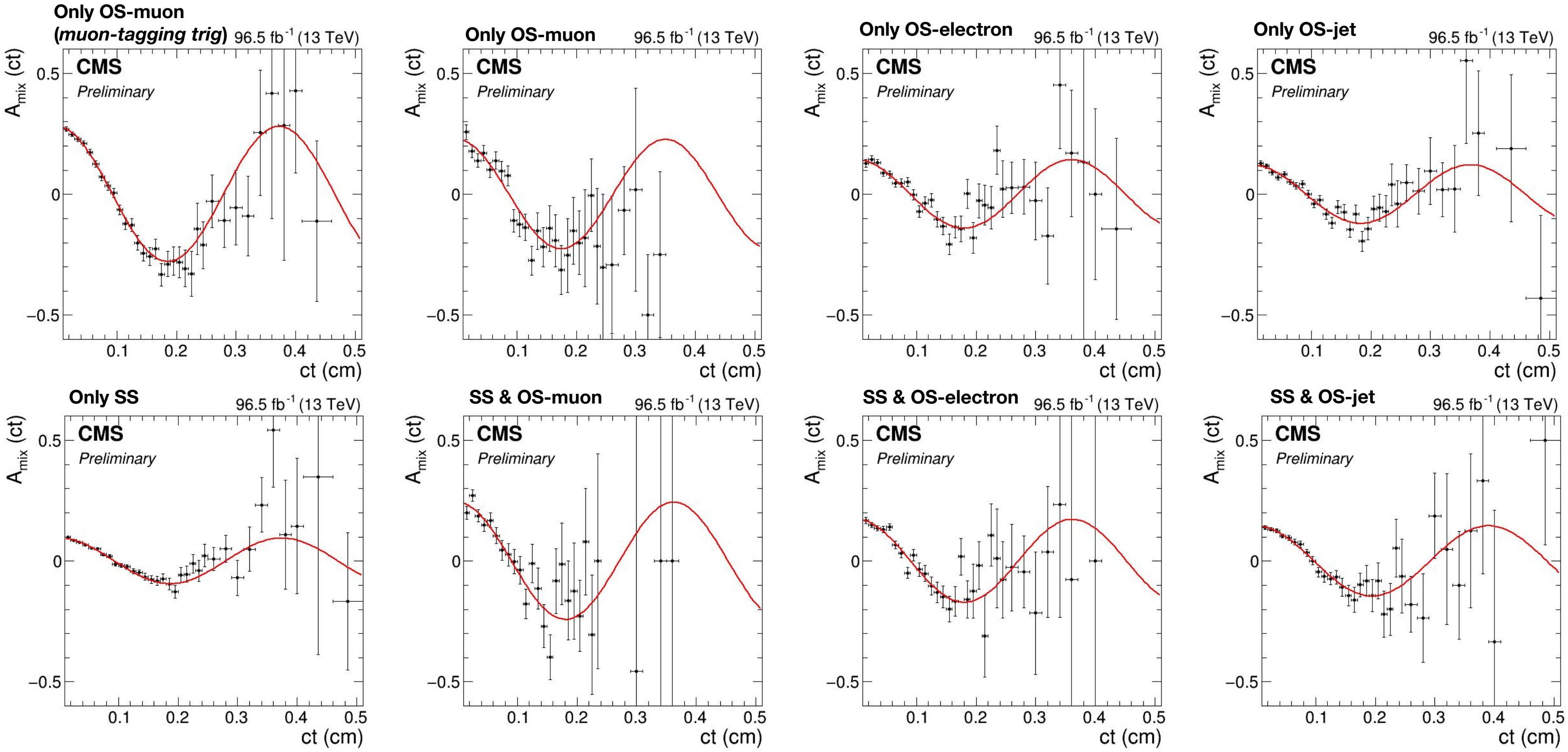
$$p(\bar{b}) = \prod_{i=1}^{2} \left(\frac{1 - \xi_i}{2} + \xi_i (1 - \omega_i) \right) \qquad p(b) = \prod_{i=1}^{2} \left(\frac{1 + \xi_i}{2} - \xi_i (1 - \omega_i) \right)$$

$$P(\bar{b}) = \frac{p(\bar{b})}{p(\bar{b}) + p(b)} \qquad P(b) = \frac{p(b)}{p(\bar{b}) + p(b)}$$

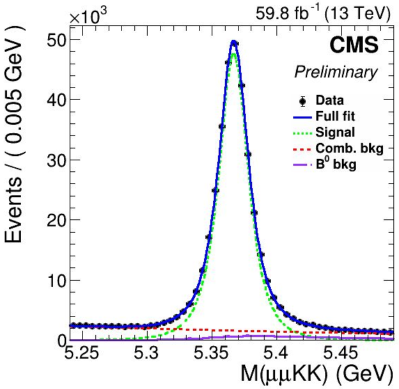
Mixing asymmetry for different tagging categories

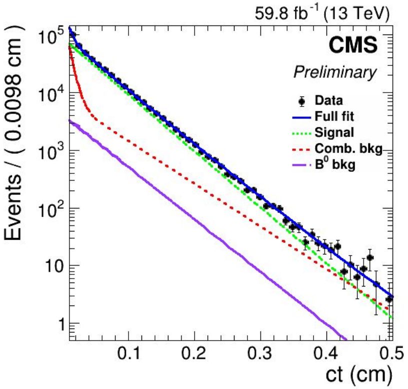
All categories are mutually exclusive

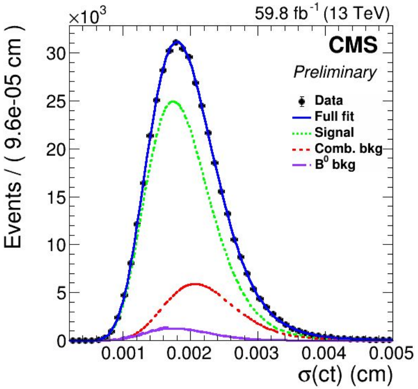
All, but the first,refers to the *standard* trigger category

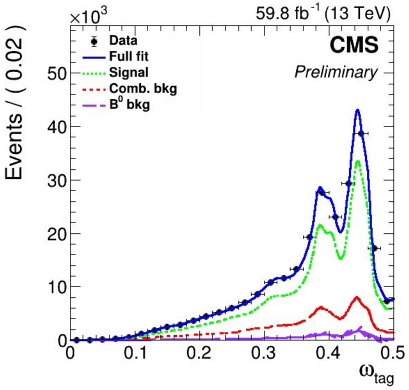


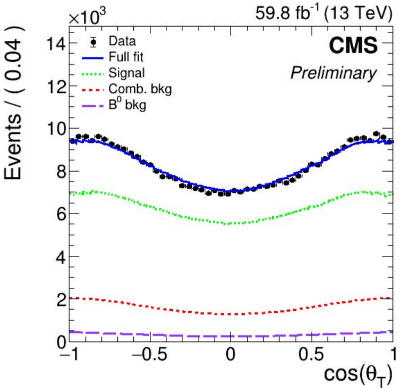
Fit projections (*standard* trigger category 2018)

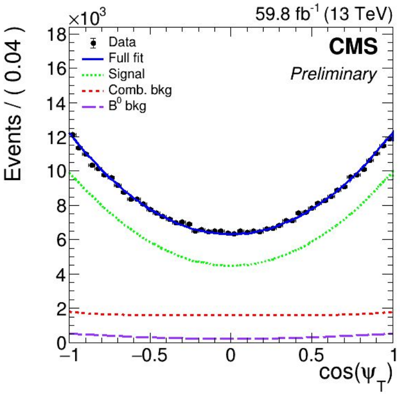


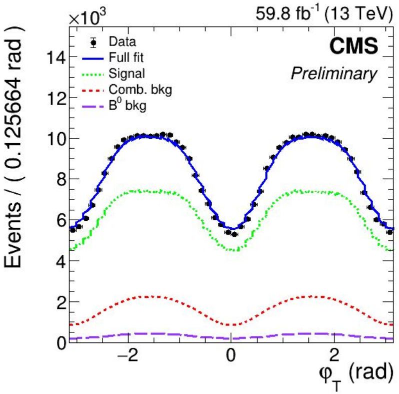












Systematic uncertainty classification

- **Type-I: unaccounted uncertainties**
  - Account for the finite statistics of simulated/control samples and uncertainties in calibrations and efficiency
  - **Always** propagated to the final results
  - Evaluated with two procedures
    1. *Type-I full*: obtained by sampling the samples/parameters of interest ~100 times, repeating the fit each time, and taking the RMS of the results as uncertainty
    2. *Type-I simple:* obtained by sampling a parameter only two times at $\pm 1\sigma_{stat}$
- **Type-II: method and model assumptions**
  - Account for **possible** bias induced by the assumptions made in the fit model and the analysis methods
  - Evaluated only if a **significant** bias is observed while testing an alternative (good) hypothesis
  - A significant bias for a parameter V is defined as a difference Δ in the fit results of **more than 20% of its σ_{stat}**
  - In these cases, **half of the bias** is taken as uncertainty, assuming that the *true* bias is uniformly distributed between 0 and Δ

The fit bias does not fall into either of these two categories

Comparison with theory and world averages

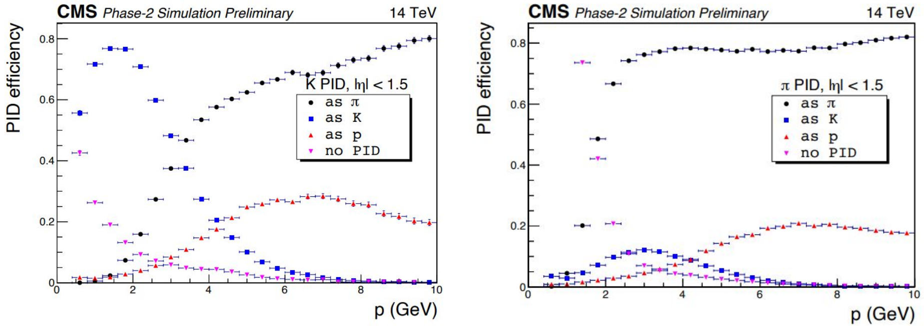
Parameter	Measured value	World-average value	Theory prediction	
ϕ_s [mrad]	-73 ± 24	-49 ± 19	-37 ± 1	[CKMfitter, UTfit]
$\Delta\Gamma_s$ [ps^{-1}]	0.0761 ± 0.0047	0.084 ± 0.005	0.091 ± 0.013	[Lenz & Tetlalmatzi-Xolocotzi]
Γ_s [ps^{-1}]	0.6613 ± 0.0032	0.6573 ± 0.0023	–	
Δm_s [$\hbar$ps^{-1}]	17.757 ± 0.039	17.765 ± 0.006	18.77 ± 0.86	[Lenz & Tetlalmatzi-Xolocotzi]
$\lvert\lambda\rvert$	1.011 ± 0.018	1.001 ± 0.018	1	
$\lvert A_0\rvert^2$	0.5300 ± 0.0047	0.520 ± 0.003	–	
$\lvert A_\perp\rvert^2$	0.2409 ± 0.0037	0.253 ± 0.006	–	
$\lvert A_S\rvert^2$	0.0067 ± 0.0034	0.030 ± 0.005	–	
$\delta_\parallel$	3.145 ± 0.078	3.18 ± 0.06	–	
$\delta_\perp$	2.931 ± 0.102	3.08 ± 0.12	–	
$\delta_{S\perp}$	0.48 ± 0.16	0.23 ± 0.05	–	

Reference: CERN-CMS-DP-2022-025

Flavor tagging in Phase-2 with MTD

- The MTD (Mip Timing Detector) provides time information of charged tracks at its surface
- The reconstruction algorithm utilizes compatible times of tracks from a vertex to offer time-of-flight based particle identification (PID) as a natural byproduct
- Same-side tagging could utilize charge correlation between the *s*-quark in the B_s and a nearby soft kaon for flavor tagging
- The PID from MTD, when integrated in the Phase-2 extrapolation of this analysis, shows a significant improvement of the tagging performances

Simulated PID efficiencies



Relative gain in P_{tag} (only SS)

PID scenario	Gains in P_{tag}
MC truth (perfect PID < 3 GeV)	+66%
PID with σ_{BTL} = 40 ps	+24%
PID with σ_{BTL} = 70 ps	+14%